DOCUMENT RESUME

ED 360 371 TM 020 275

AUTHOR Ingels, Steven J.; Scott, Leslie A.
TITLE Exclusion of Students with Barriers to Participation in NELS:88--Baseline Excluded Students Two and Four Years Later.
INSTITUTION National Opinion Research Center, Chicago, Ill.
SPONS AGENCY National Center for Education Statistics (ED), Washington, DC.
PUB DATE Apr 93
CONTRACT 300-86-0010; RS8800-30001; RS90005001
NOTE 19p.; Paper presented at the Annual Meeting of the American Educational Research Association (Atlanta, GA, April 12-16, 1993).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Attendance; *Data Collection; Eligibility; *Followup Studies; Grade 8; Junior High Schools; Junior High School Students; Limited English Speaking; *Longitudinal Studies; National Surveys; Research Methodology; *Research Problems; Sample Size; *Sampling; *Test Bias
IDENTIFIERS *National Education Longitudinal Study 1988

ABSTRACT

Sample undercoverage issues in the National Education Longitudinal Study of 1988 (NELS:88) are addressed. The main focus is the exclusion of certain categories of student in the base year, 1988, and in in-school follow-up rounds. A subsidiary focus is the question of how adequately transfer students were captured within the sampling procedures of the study. Recommendations are offered for how better to deal with undercoverage issues in future school-based longitudinal studies. The six ways in which a student might not have been selected were: (1) refusal by the school to participate; (2) ineligibility of the school; (3) ineligibility of the student, for language, disability, behavioral problems, or lack of English; (4) absence from the school due to study elsewhere; (5) temporary unavailability due to illness or transition; (6) clerical error; and (7) inadequate sampling frame that omitted a school. The exclusion of students is referred to as a problem, but including everyone would have been more of a problem. Ways to increase the rate of meaningful participation in the future are discussed. The experience of NELS:88 suggests that more students have been excluded than is justified. Two tables provide study data. (Contains 28 references.) (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

STEVEN J. INGELS

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED360371

Exclusion of Students With Barriers to Participation in NELS:88-- Baseline Excluded Students Two and Four Years Later.

Steven J. Ingels, Leslie A. Scott

National Opinion Research Center
University of Chicago
1155 E. 60th Street
Chicago, Illinois 60637-2799

Paper presented at the American Educational Research Association Annual Meeting, Atlanta, Georgia, Session 25.60, April 14, 1993.

This paper is based on the National Education Longitudinal Study of 1988, sponsored by the National Center for Education Statistics (NCES), Office of Educational Research and Improvement, U.S. Department of Education. Dr. Jeffrey A. Owings was the NCES Project Officer for the NELS:88 Base Year; Dr. Shi-Chang Wu is the NCES Project Officer for the first follow-up; Peggy Quinn is the NCES Project Officer for the second follow-up. The ideas and opinions expressed in this paper are those of the authors and do not necessarily represent positions or conclusions endorsed by NCES.

NORC Projects 4455, 4492, 4521; NCES Contracts 300-86-0010; RS8800-30001; and RS90005001.

TM 020275

BEST COPY AVAILABLE

This paper addresses sample undercoverage issues in a national education longitudinal study, the National Education Longitudinal Study of 1988 (NELS:88). Its main focus is the exclusion of certain categories of students in the base year of the study and in in-school follow-up rounds. A subsidiary focus is the question of how adequately transfer students were captured within the sampling procedures of the study. Recommendations are offered for how better to deal with undercoverage issues in future school-based longitudinal studies.

1. *Ineligible and Excluded Students*

There are a number of possible sources of undercoverage bias. In particular, there are seven ways in which a student may have failed to have a chance of selection into the NELS:88 base year.[1]

(a) First, if the student's school refused, that student had no chance of selection;[2]

(b) Second, if the student's school was declared ineligible to participate, that student had no chance of selection;

(c) Third, though the selected school participated, the student was declared ineligible to participate, owing to physical or mental handicaps, behavioral problems, or a lack of command of English;

(d) Fourth, the student was studying at home in 1987-88, or abroad, or in an ungraded program or school;

(e) Fifth, the student was temporarily unavailable (for example, was hospitalized during the survey period, or was a migrant in transit);

(f) Sixth, owing to clerical error, the student did not appear on the correct roster or was misclassified. (While we believe that in general school rosters were extremely accurate, there is some evidence that transfer-ins between the time of initial sampling and the sample update just before Survey Day were, as a group, somewhat underrepresented);

(g) Seventh, the student's school had no chance of selection, because the sampling frame was inaccurate (for example, a student might attend a newly-opened school that had not yet been added to the school list from which the sample was drawn).

---

[1]Through the process of sample freshening, 1990 sophomores and 1992 seniors who had no chance of selection into NELS:88 in 1988 (because they were not in the United States, or not in the eighth grade), are added to the dataset. Most (but not all) of the same sources of potential undercoverage arise in freshening as in base year sample selection. As an example of the exceptions to this generalization, one is no longer dependent on a published universe list of schools and therefore vulnerable to omissions in the sampling frame, in that the schools at which freshening took place in 1990 and 1992 were the schools to which 1988 eighth graders dispersed, nor was any type of school ineligible in the follow-ups.

[2]Substitute selections replaced original selections that refused. Potential school nonresponse bias is analyzed in Spencer, Frankel, Ingels, Rasinski & Tourangeau, 1990.

The focal point of this paper is category (c), excluded students, although we will make some observations as well on categories (b), (d), (e), and (f). We will comment briefly on these categories below.

Ineligible schools (b). Virtually all schools in the fifty states and the District of Columbia that enrolled eighth graders in the 1987-88 school year were eligible for the study. However, Bureau of Indian Affairs (BIA) schools were categorically excluded from the 1987-88 school frame. Given that only about 1 percent of eighth graders at the time were American Indians and that 90 percent of American Indian students attend non-BIA schools, this exclusion should have a negligible impact on estimates, though it should be taken into account when considering results for the American Indian subgroup. Also excluded were special education schools for the handicapped, area vocational schools that do not enroll students directly, and schools for dependents of U.S. personnel overseas. Students at schools that do not enroll students directly presumably had a chance of selection into NELS:88 through the schools in which they were directly enrolled. Students outside the fifty states and the District of Columbia were defined as out of scope on the supposition that they were not of interest in a study of schooling in the United States. Insofar as special education schools often utilize ungraded programs, they do not fit the grade cohort definition of eligibility on which the NELS:88 student sample was built. To the extent that grade designations may be in use, the presumption was made that students who are in specialized schools will tend to be more severely handicapped than those who are mainstreamed, and would not readily be able to complete the requirements of the testing program.

Excluded Students (c). The excluded students are a subclass of the ineligible students, specifically, those who were declared ineligible for reasons of mental, physical, or linguistic barriers to participation.[3] While students who died, were part-time students primarily registered at another school, or who transferred out of the school prior to its Survey Day, were also declared ineligible, these categories of students should affect neither the representativeness of the sample nor estimates derived from it. The governing principle here is that each 1987-88 eighth grader should have one chance of selection into the NELS:88 sample, and only one. Part-time students with a primary registration elsewhere had a chance of selection into the sample at the site of their primary registration. Transfers out of the school were classified as ineligible, but sample representativeness was maintained by giving transfers into the school during the same period a chance of selection into the Base Year sample. However, the 5.4 percent of base year students with severe physical, mental or linguistic obstacles to participation had, as a class, no chance of selection into the sample; they were systematically excluded.[4] Assuming their characteristics and behaviors to be in any essential way different from the norm, their exclusion will be a source of undercoverage bias in national estimates.

Home Study, Abroad, Ungraded (d). While students not enrolled in an American school but receiving an education at home or abroad were not eligible for selection into the base year, such students had a chance of selection into the study in 1990 or 1992, if their status had changed, that is,

---

[3]For further details see the NELS:88 base year sample design report: Spencer, Frankel, Ingels, Rasinski & Tourangeau (1990); also Ingels, Rizzo and Rasinski (1989) and Ingels (1991).

[4]Total eighth grade enrollment for the NELS:88 baseline sample was 203,002 students, of whom 10,853 were excluded owing to limitations in their language proficiency (35 percent of the exclusions), physical disability (8 percent of the exclusions), or mental disabilities (57 percent of the exclusions).

if they were in the tenth grade in a school in the United States in the 1990-91 school year or the twelfth grade in an American school in the 1991-1992 school year. Implicitly, students in ungraded programs (which, historically, have been fairly common for students with severe handicaps[5]) are excluded, since NELS:88 is a grade cohort, not an age cohort, and such students will not appear on an eighth grade school roster.

Temporarily unavailable (e). Students undergoing prolonged hospitalization or institutionalized or otherwise unavailable were extremely rare in the base year. However, in the NELS:88 follow-ups, substantial numbers of students, particularly in the northeast and on the west coast, had left the country at the time of data collection. Such students are regarded as temporarily out of scope in NELS:88, and subject to re-survey should they have returned to the United States at the time of the next data collection. Migrant students may be a group that is particularly hard to represent within a school-based sample. Generally the most stable period for sampling this group, that is, the time at which they are likely to be at their home base school and not in transit, is early in the calendar year. Because of the small size of the migrant student population (about seven tenths of one percent of public school enrollment, per Henderson, Daft and Gutmann, 1989), some under-representation of this group would not pose a large risk of biasing national or subnational estimates.

Misclassifications and omissions (f). A small number of cases have been removed from the NELS:88 sample owing to later discovery that the student was in a grade other than grade 8 at the time of sampling, and appeared on an eighth grade roster in error. Presumably some number of cases that should have been listed on eighth grade rosters did not appear. While the number of such cases is likely to be quite small, there is no way to be certain precisely how many eighth graders may have been omitted from school listings.

However, undercoverage of **transfer** students is a quantifiable problem. NELS:88 followed essentially the same procedure for dealing with transfer students as did High School and Beyond (HS&B) in 1980. School rosters were submitted and an initial sample drawn in the autumn. To adjust the student sampling frame for student attrition and change in the eighth grade population of the sampled school, NORC conducted a sample update seven to ten days prior to the school's scheduled survey session. The NORC survey representative went over the sample list with the school coordinator to ensure that all sampled students were still eligible, and that transfers-in--that is, any student who had joined the eighth grade class of the school between the time of the original sampling and the time of the update--were added to a supplementary roster from which additional students would be selected. Selections for inclusion in the sample were based on the same set of computer-generated random numbers used to select the original sample.

Given that mortality and dropout rates are very low within eighth grade cohorts, in theory, there should be a rough parity in the number of selected students lost to transfer and the number selected into the sample from the pool of transfers in. Overall, around four percent of the NELS:88 original sample had transferred out by survey day, but the replacement rate was around two percent, half the expected percentage (Ingels, Rizzo and Rasinski, 1989). This experience is not peculiar to NELS:88. For example, for the National Assessment of Educational Progress (NAEP) Trial State Assessment in 1990, Spencer (1991b, p.6) reports that 4.9 percent of the students withdrew from the

[5]Nevertheless, at present "91 percent of elementary and secondary public special education students are in graded classes (or placements)" according to the National Council on Disability (1993).

sample but the supplemental sampling procedure added only 2.9 percent to the sample. Thus there was a 40 percent undercoverage of transfer students. The reason for undercoverage of transfers in the base year of a longitudinal study or in a cross-sectional study would appear to be that while all transfers out will be identified successfully (any missed outward transfers at the sample update will be identified when the no-shows at survey day are investigated), school records are not always sufficiently accurate and up to date to provide definitive lists of all students who have transferred in since a certain date.

In the NELS:88 case, it should be noted that, in principle, transfer students are followed in the longitudinal follow-ups, given that the sample is frozen with the base year survey session. In practice, it has not been possible to follow all transfer students, given the enormous dispersal between grades eight and ten. Hence smaller student clusters were subsampled in the first follow-up. Even after sample selection in 1989, students continued to transfer at a high rate, so that, for cost reasons, a 20 percent subsample was taken of transfer students in 1990.[6] While sample weighting permits transfer students to be properly represented in the study, numerically, the sample of transfer students is comparatively thin. Nor was any measure taken to compensate for the undercoverage of transfers in the base year.

2. ***Statistical and Equity Problems Associated with Excluding Students from the Sample.***

Of the various sources of potential undercoverage bias that we have discussed, the exclusion of students for reason of linguistic, mental or physical barriers poses the greatest threat to the task of providing reliable national estimates from survey data. Excluded students are a problem for multiple reasons. Excluded students may be a source of bias in national estimates both overall and in particular for certain policy (e.g., IEPs, LEPs [i.e., special education students with an Individualized Education Program, or English non-proficient or limited-proficient students]) or demographic (e.g., Asians, Hispanics) groups. Any undercoverage bias introduced in the Base Year will persist in subsequent rounds of the study. Additionally, if any of the reasons for exclusion are based on individual traits that may change over time, the representativeness of the tenth and twelfth grade samples will be compromised if 1987-88 eighth graders who have overcome their barrier to participation in the meantime are given no chance of reselection into the study.

The potential impact of exclusions on national estimates can be readily illustrated. The phenomenon of dropping out offers one of several possible examples. One of the most important of the purposes of NELS:88 is to investigate the dynamics of school leaving and school completion. The probable understatement of dropout rates and loss of representativeness to the dropout sample attendant upon these exclusions must be taken as a serious consequence of having incomplete (95 percent) representation in NELS:88 of the 1987-88 eighth grade population. Subgroup estimates may be particularly affected. For example, exclusion owing to language barrier may particularly affect groups with high recent immigration rates to the United States, including many Hispanic and Asian

[6]For details, see Chapter 3 of Ingels, Scott, Lindmark, Frankel, and Myers, 1992.

subgroups.[7] Additionally, excluded members of the subgroup are likely to differ in other respects also from included members. Moreover, if a high proportion of language minority and handicapped students are excluded, capacity to study these highly policy-relevant subgroups will be severely diminished. Later in this paper, we will compare national dropout rates derived from consideration of the eligible-only sample and the eligible and ineligible sample, to measure the impact of ineligibility rules on national dropout estimation.

It is sometimes maintained that there are non-statistical problems with excluding students from national research studies--and, in particular, state and national assessments--as well. McGrew, Thurlow, Shriner and Spiegel (1992) argue that exclusion of students with disabilities from major research and testing programs is a problem from an equity perspective. Indeed, assessments in particular are commonly seen as critical agents of reform, which may affect the motivation of students, the content of the curriculum, and the skills and techniques of teachers (Linn, 1993). It is in the context of such considerations that some have claimed (National Council on Disability, 1993) that exclusion of substantial numbers of disabled students from national data bases that contain achievement measures has situated such students at the periphery of recent school reform. Assessment-driven reform also of course raises the question of "what to measure" and whether the developmental needs of all students are currently being taken into account. Thus Bruininks, Thurlow and Ysseldyke (1992) point to the need to refine the models of educational outcomes that are critical to enhancing the cultural assimilation and quality of life of handicapped students so that proper indicators may be put in place as part of the overall assessment system..

3. *How Well Were the Base Year Exclusion Criteria Applied?*

While extreme handicaps are readily identified, there is oftentimes ambiguity about other disability conditions. Owings and Stocking (1985) examined student self report data on handicap status in HS&B, and found that reports of handicaps were not stable over time. For many students who reported less severe disabilities, self-perceived handicapped status was a condition dependent on various factors and subject to change, not an enduring trait. Nonetheless, student reports of handicap status were systematically related to other characteristics (for example, lower self-esteem), and can be seen as pointing to the importance of attention to the special needs of self-identified handicapped students.

If there are difficulties with subjective classifications of disability, classifications by school personnel also appear to be less reliable than one would wish (Bennett and Ragosta, 1984). Moreover, certain racial/ethnic minorities have disproportionately been placed in special education classes (National Council on Disability, 1993, pp. 45-49). To take the most prevalent handicapped status--learning disability--as an example of the reliability problem, Bennett and Ragosta (1988) observe (p.19):

---

[7]If anything, the problem of language barriers to test-taking and survey participation is likely to increase in coming years, if current immigration trends continue; the degree to which native language instruction is used in American schools may also influence the extent of language-based exclusions. Decennial census data (Bureau of the Census, 1993) shows that in 1990 31.8 million Americans over the age of five spoke a foreign language at home, compared to 23.1 million in 1980 (however, three quarters also spoke English, either "well" or "very well"). Spanish-speakers in U.S. residence increased 50 percent over the 1980s, to 17.3 million in 1990.

Because current federal definitions of *learning disability* give little practical guidance as to what it is, some 40 competing definitions have been suggested (Ysseldyke, 1983). The breadth of characteristics encompassed by these definitions is so great that a study of 17 of them found 85 percent of a normal student sample classifiable as learning disabled (Ysseldyke, Algozzine, & Epps, 1983).

Bennett and Ragosta go on to summarize the structure and implementation of the special education classification and placement process (1988, p.20):

> Typically, students enter the process as a result of referral by their classroom teachers. Research has shown that such referrals are sometimes based on such extraneous factors as race, sex, physical appearance, and socioeconomic status....Assessment is followed by a classification meeting at which a diagnosis is made. Investigations of this aspect of the process report little consistency in diagnostic statements among professionals assigned to the same case, only a slight relationship between assessment data and team judgments, and the influence of irrelevant pupil characteristics on classification decisions....As a result, studies suggest that over half of the classification decisions made by child-study teams are erroneous (Algozzine & Ysseldyke, 1982; Craig, Myers, & Wujek, 1982; Shepard, Smith, & Vojir, 1983). One effect of these placement errors is to confuse attempts to characterize the true nature of handicapping conditions even further.

It is important to keep in mind that NELS:88 excluded students were determined by their schools to be unable to participate. Criteria for exclusion were provided to the schools, but it was up to the school itself--usually the School Coordinator or the principal--to interpret and apply the eligibility criteria. It is also important to note that schools were asked to apply the criteria on an individual basis. Thus, limited English proficient (LEP) students or special education students were not to be excluded categorically. Rather, only those particular LEP or special education students whose limitations were so severe as to constitute significant barriers to meaningful participation were to be excluded. In cases of uncertainty, school personnel were asked to include the student.

In general, this process worked reasonably well in that almost certainly the extreme cases of physical or mental disability, and limitation of English proficiency, were excluded. A few students were included who manifestly should not have been. Their difficulty in completing the questionnaires and tests was noted by survey administrators, and Educational Testing Service rejected as unusable a small number (less than one percent)[8] of cognitive tests.

---

[8]Completion rates were in excess of 99 percent for all tests. Sections were not scored if fewer than five items were answered in the section; most students in this group answered no items at all. Then a "reasonableness check" was performed to identify students with ten or fewer items answered and whose IRT-estimated scores were more than three points higher than their raw scores. (This can happen if, for example, a student answers the first seven items correctly and then, owing to lack of motivation or some other factor, fails to complete the rest of the test. The IRT-estimated score would be the highest possible score on the test since there were no wrong answers), but the formula score would be only seven. Since the reason for the discrepancy was unknown, these scores were deleted.) Most deleted cases had zero items answered, and some of these cases could represent students who found the tests too difficult to attempt. The percentage of usable cases was 99.7 percent in reading and mathematics, 99.5 percent in science, and 99.2 percent in history/geography/citizenship. (IRT [Item Response Theory] is a method of estimating ability level by considering the pattern of right, wrong, and omitted responses on all items administered to an individual student.)

However, the fact that such cases were rare suggests that the exclusion procedure was in the main effective in screening those students who should not have been asked to participate. Indeed, one could draw the conclusion that the screening out of students was too effective in that one would expect more such cases had schools taken with full seriousness the injunction "when in doubt, include."

In any case of the application of general criteria, there is bound to be some degree of arbitrariness in judgments about borderline cases. This arbitrariness is of course compounded when the numbers of people rendering judgments about such marginal cases is large, as was so in the NELS:88 Base Year. Our largest concern about the classification process, however, is that, for reasons of time and burden, some schools (perhaps as many as a hundred of the 1,052 schools in the sample) apparently departed from their instructions and excluded students on a categorical basis in preference to rendering the prescribed case-by-case assessments. (Evidence for this phenomenon is seen when sampling rosters are inspected and all students within a pre-existing category are excluded.) In consequence of categorical exclusion, and in consequence of non-categorical exclusions based on minimal information for evaluation, one would expect that overall, more students may have been excluded than necessary. The temptation to exclude categorically--in a school with a large eighth grade, given severe time pressures for producing an annotated roster, and with individual-level information available to the School Coordinator only through the laborious process of interviewing the special education or bilingual education teacher of each student--is large.

In order to minimize this problem in the future, greater precision in exclusionary definitions should be sought. Setting out more specific conditions for ineligibility would increase school burden and might adversely affect prospects for cooperation in a few cases, but in general would maximize the number of participants by minimizing the number of wrongful exclusions. (One might, for example, achieve greater definitional specificity by saying that to exclude a student for mental or physical disability, that student should normally be a special education student with an Individualized Education Plan [IEP] who [a] **is not mainstreamed in English/language arts**; and [b] is judged by the school to not be capable of completing the survey forms. Invocation of an IEP and decision rules based on mainstreaming is a tack taken by NAEP, starting with the 1990 assessment[9]). In addition, giving schools the option of excluding students from the test while including them for purposes of questionnaire administration would further minimize any problems of excessive exclusion.

4. *1988 Base Year Excluded Students in 1990 and 1992*

NORC selected from the greater pool (over 10,000) the number of excluded students who would have been included in the study had no students been excluded. That first cut yields 1598 base year excluded students. Had there been no exclusion, the NELS:88 base year sample (N = 26,432) would have contained an additional 59 students with physical disabilities, an additional 835 students with mental disabilities, and an additional 532 students with language barriers to

---

[9]Of course, there is great local variation in the operationalization of the concept of an IEP (and so too for LEP status). Diagnoses of disabilities (including emotional disturbance and learning disability) are not highly reliable, and schools may have funding and other incentives for using these labels. Moreover, it is claimed that 9 percent of students with true disabilities either do not have an IEP or have not been properly evaluated (National Council on Disability, 1993).

participation.[10] In addition, 172 students were excluded with no reason given for their exclusion. This sample of 1598 base year excluded students was subsampled for budget reasons; a subsample of 674 base year ineligibles were pursued in the first follow-up (1990) and second follow-up (1992). Of the 674 excluded base year students, 225 were language exclusions, 24 were physical disability exclusions, 352 were mental barrier exclusions, and 73 were excluded with no reason given.

First Follow-Up Results: **overall results**. Of the 674 base year excluded students studied in the first follow-up, NORC was able to ascertain the status of all but 42. Hence information on school enrollment status and NELS:88 eligibility status was obtained for 94 percent of the excluded student sample. Some 48 exclusions were found to be sampling errors (for example, the student's name appeared on an eighth grade roster, but the student was not an eighth grader, owing to retention in the prior grade or some other factor; or the student's name appeared on the school's roster but the student had transferred out or had never enrolled). Removing these 48 cases provides a new sample size of 674 - 48, or 626.

Of the 626 cases, 29 were declared out of scope, because of either the death of the sample member, or the sample member being outside the country in the spring term of 1990 (such cases are viewed as only temporarily out of scope--such individuals would be pursued in 1992 in cases where they had returned to the United States). If these cases are subtracted from the denominator, a sample size of 597 is obtained. Of those 597 students, 314 were found to be eligible, 241 were found to be still ineligible, and the status of 42 was not ascertained. In other words, of the 597 in scope base year excluded students in 1990, the enrollment and eligibility status of 7 percent could not be ascertained (mostly, these cases were unlocatable), 53 percent were found to be eligible for NELS:88, and 40 percent were still ineligible.[11]

First Follow-Up results: language exclusions. These results can be viewed for each of the categories of exclusion, thus language, physical, and mental barriers to participation. For language exclusions, almost 72 percent (131) of in-scope respondents were reclassified as eligible, nearly 22 percent (40) retained their ineligible classification, and around 7 percent were unlocatable and their status could not be ascertained.

First Follow-Up results: physical handicap exclusions. Of 23 physical barrier exclusions, 39 percent (9) were reclassified as eligible in 1990, 52 percent (12) remained ineligible, and about 9 percent (2) could not be located.

First Follow-Up results: mental handicap exclusions. Of 333 in-scope base year ineligibles excluded in 1988 by virtue of mental barriers to participation, 42 percent (140) were classified as eligible in 1990, almost 53 percent (175) as ineligible, while for 5 percent (18), status could not be ascertained.

---

[10]While all excluded students were assigned to one of the three exclusion categories (or to the "no reason" category), in a handful of cases students had multiple bases for exclusion (for example, one might have both a physical and a language barrier to participation).

[11]All percents are raw (sample) percents; weighted percents, which supply national population estimates, could differ.

Results at the Conclusion of the Second Follow-Up. Second follow-up results show (see the two tables below) comparatively little change from the first follow-up, though the numbers of ineligible students have diminished further, with the proportion eligible increasing from 53 percent to 57 percent.

Table 1: Summary of Final 1992 Statuses for 1988 Excluded Students in Unweighted Percents

Reason for 1988 exclusion:	ELIGIBLE	INELIGIBLE	NOT ASCERTAINED	SAMPLE N
language barrier	70.6%	12.4%	16.9%	177
physical barrier	56.5%	39.1%	4.3%	23
mental impairment	50.2%	42.3%	7.6%	331
unknown reason	54.5%	27.3%	18.2%	55
TOTAL	57.0%	31.7%	11.3%	586

(excludes cases sampled in error and those out of scope [dead or out of country] for 1992 round)
(owing to rounding, rows may not sum to 100 percent)

Table 2: 1992 Status Ns of 1988 Excluded Students

1988 reason for exclusion:	ELIG.	INELIG.	OUT OF SCOPE	N.A.[12]	SAMPLING ERROR
language	125	22	25	30	23
physical	13	9	0	1	1
mental	166	140	5	25	16
unknown	30	15	2	10	16
TOTAL	334	186	32	66	56

[12] N.A. = status not ascertained.

One other perspective for examining the status of base year excluded students two years later is to note the impact of this small group on national dropout estimates.

Impact of Excluded Students on National Dropout Estimates. It is generally recognized that while many physically impaired students may have low dropout rates and may indeed be eligible for an atypically long period of public education, other groups, such as those with mental impairments or limited English proficiency, drop out of school at disproportionately high rates.[13] Below, we compare national dropout estimates for the eighth grade cohort between 1988 and 1990 that reflect inclusion and exclusion of the 5.4 percent of the base year sample that was ineligible in 1988. (Dropout rates for the second follow-up [1988-1992, 1990-92] have not yet been calculated as of this date.)

Table 3: Eighth Grade Cohort Dropout Rate, 1988-1990.
(Percentage of Spring Term 1988 Eighth Graders Not in School Spring Term 1990)

	ELIGIBLE SAMPLE		EXPANDED SAMPLE	
Total.	**6.05 %**	(0.48)	**6.82 %**	(0.40)
Race/Ethnicity.				
Asian	3.1	(1.05)	4.0	(1.02)
Hispanic	9.2	(1.01)	9.6	(0.84)
Black	10.0	(1.94)	10.2	(1.51)
White	4.9	(0.53)	5.2	(0.44)
American Indian	10.5	(2.60)	9.2	(2.32)
Gender.				
Male	6.3	(0.69)	7.2	(0.55)
Female	5.8	(0.59)	6.5	(0.51)
1988 Eighth Grade Public School Students.	6.8	(0.55)	7.6	(0.45)

Note: standard errors appear parenthetically after each estimate.

Source: National Education Longitudinal Study of 1988 (NELS:88) First Follow-Up, National Center for Education Statistics.

---

[13]The National Council on Disability (1993, pp.8-9) indicates that "Students with serious emotional disturbances are at the greatest risk among all student groups of dropping out of school, at a rate of about 40%". (See also Wolman, Bruininks and Thurlow, 1989).

Interpretation of 1990 and 1992 Results. Because of their disproportionately high overall dropout rate, inclusion of excluded students in dropout statistics does matter to overall estimates. Moreover, reassessment of eligibility status led to inclusion in NELS:88 follow-up rounds of a majority (57 percent) of the students found ineligible in the base year. Of those excluded, those excluded for language reasons had the greatest chance of re-entering NELS:88 by 1992, with 72 percent eligible, 12 percent ineligible, and 17 percent unlocatable four years after their exclusion in eighth grade.

In contrast, about 57 percent of the small physically ineligible sample, and about half of the mentally impaired sample, had been reclassified as eligible by 1992. It is unsurprising that such a high proportion of the "unknown" category (no reason for exclusion given) turned out to represent sampling errors (about 22 percent of this group appeared on rosters in error).

These changes in status represent several tendencies that cannot readily be disentangled. First some students' status will have changed. This result is most likely for English non-proficient and limited proficient students, who over time may master English to a significantly greater degree. Second, judgments of ineligibility, though guided by objective criteria, also have a subjective dimension, and are somewhat unreliable. Some amount of change will be associated simply with re-asking the eligibility status question. Third, the question of eligibility was not posed in precisely the same way in 1990 and 1992 as in the 1988 base year. Though the general criteria were largely unchanged[14], further information was provided for the interpretation of the general criteria.[15] In addition, information was sought from school staff who had a greater likelihood of personally knowing the student. The task, for school personnel, of supplying information about a small number of base year ineligibles was far less daunting and presumably less error-prone than the task of providing classification information for up to several hundred potential sample members per school in the base year. Still, in the main, these considerations point to the likelihood that the 1990 and 1992 classifications are more accurate than the 1988 classifications, in instances where the individual has not significantly changed, and the likelihood that where change has occurred in a student's eligibility status, that change has been captured. These considerations also support the contention that a large number of students who could successfully have participated were excluded by their schools.

Nevertheless, in the conditions of mounting a new large-scale, multi-level, high-burden survey, it must be recognized that incentives to over-exclude would again be present. (In the NELS:88 base year, for example, it was necessary for 1,052 school coordinators to screen over 203,000 eighth grade students for eligibility status; over 10 percent of schools had eighth grades of 400 or larger; for a study with a starting point in high school, class sizes tend to be much larger still). It must be recognized as well that a substantial pool of ineligibles (defined as those unable to complete the survey forms) will remain, even if the numbers of excluded students are substantially reduced. Even after triple screening and the passage of four years during which some individuals became more proficient in English or underwent other status changes, about a third of the 1988 NELS:88 ineligibles were still ineligible.

---

[14]A change that affected a very few Hispanic ineligibles was the provision of a Spanish-language NELS:88 questionnaire in 1990, and again in 1992; a Spanish language student questionnaire was not offered in the base year.

[15]For further documentation of screening procedures, see Ingels 1991.

5. ***Recommendations for Future Studies***.

Possible Means for Reducing the Proportion of Excluded Students. Although we have referred to the exclusion of students from the NELS:88 sample as a problem--and a statistical sampling problem it doubtless is--it would of course be even more of a problem to include everyone. The principle of exclusion recognizes the simple fact that not everyone is capable of participating in such a study. It would be ethically unconscionable, and futile as an exercise in data collection, to have a non-English proficient student, or an educable mentally retarded student, struggle for eighty-five minutes to attempt to complete a cognitive test that the student simply could not comprehend. It would be imprudent to ask a student who has behavior control problems and cannot concentrate for sustained periods to attend a three-hour survey session with other students. Nor could schools be expected to cooperate with a study that made such demands. Schools are often (and quite rightly) unwilling (and sometimes legally constrained) in the matter of allowing students with certain handicaps or certified low reading levels to participate.

For these reasons, High School and Beyond (HS&B), NELS:88, and the National Assessment of Educational Progress (NAEP) have all had policies of excluding certain categories of the school-age population.[16] Indeed, McGrew, Thurlow, Shriner and Spiegel (1992) in their study of 9 major national data collection programs, estimate that 40 to 50 percent of school-age students with disabilities are excluded from these national data collection programs. (For the NELS:88 base year, a similar proportion of the LEP population was excluded as well.)

In general the excluded student problem is far less acute for studies starting in tenth grade or twelfth grades than for studies such as NELS:88 and NAEP, that begin with or include pre-high school populations. NAEP, for example, excludes those students who are deemed by the school to be unable to participate owing to: no or severely limited English language proficiency, functional disability, or mental disability (for example, being classified as educable mentally retarded). In its 1988 assessment, NAEP excluded 5.3 percent of eighth graders, 3.7 percent of twelfth graders, and 6.3 percent of fourth graders (Johnson and Zwick, 1990). These exclusion rates show an increase over 1984 rates (for example, only 3.6 percent of eighth graders were excluded in 1984). NAEP's 1988 eighth grade exclusion rate of 5.3 percent, and the NELS:88 1988 eighth grade exclusion rate of just under 5.4 percent, are surprisingly close. Although the NELS:88 exclusion rate is not unexpectedly high, one still must ask what further measures might be employed to increase the rate of meaningful participation and to thus increase the power of survey estimates.

---

[16]An overall exclusion rate is not reported in the HS&B documentation. Hoachlander (1992, NCES 91-667) notes that "according to Harnisch, Lichtenstein, and Langford [Harnisch, D.L., Lichtenstein, S., Langford, J.B., 1986, Digest on Youth in Transition, Champaign, Illinois] 94 percent of the students who can be positively identified as handicapped in HS&B were physically handicapped; the national rate of physical disabilities among school-age children with special needs is 4 percent. Only 6 percent of the students identified as handicapped in the HS&B sample were learning disabled, and none ere emotionally disabled or retarded. The vast majority of all handicapped students is generally comprised of these three disability groups, so the sample of handicapped students in HS&B....is in no way representative of the national population of handicapped students."

The growth in performance-based[17] assessment and use of portfolios in testing repertoires may increase the opportunities to include students who have been excluded from traditional testing programs, since performance-based assessment is in principle individually modifiable. McGrew, Thurlow, Shriner and Spiegel (1992) urge the importance of including categories of students at-risk of being excluded in the instrument development process, so that instrumentation is better adapted to measuring their achievement and questionnaire responses.

Also, with additional funds, the number of excluded students in national studies could be further reduced by the following means:

* Translation of questionnaire and tests into Spanish, Chinese, Korean, and other languages associated with groups that have a high immigration rate into the United States, or otherwise constitute language minority communities (for example, American Indians).

* Extended survey administration time limits for handicapped and language minority students[18]

* One-on-one oral administration of questionnaires for poor readers and handicapped students

* One-on-one oral administration of questionnaires for NEP and LEP students, using bilingual interviewers (Chinese-English, Vietnamese-English, Spanish-English), etc.

* Large print versions of the instrument for the visually impaired or cassette/tape versions (or braille versions)

* Different assessment techniques may be appropriate for students with different impairments (for example, multiple choice tests may be less appropriate for dyslexics than essay questions or a performance-based assessment)

Useful as such measures would be, they are costly[19] and would only marginally increase the number of students surveyed. Nor are these measures wholly without difficulty in concept and application. Some who are non-proficient in English lack literacy in their mother-tongue as well; therefore production of multilingual materials does not remove the language barrier for all students in these language groups; such students also may feel other inhibitions to use of non-English forms (for whatever reason, few students in HS&B or NELS:88, for example, opted for the Spanish translation

---

[17]A performance-based test is defined by the GAO (1993) as "A test that measures ability by assessing open-ended responses or by asking a person to complete a task. Also known as alternative assessment, constructed response, or task performance, performance-based tests require the respondent to produce a response or demonstrate a skill or procedure. Examples include answering an open-ended question, conversing in a foreign language, solving a mathematics problem while showing all calculations, writing an essay on a given topic, or designing a science experiment."

[18]Extra time may increase comparability of results for handicapped and non-handicapped students, or it may diminish comparability. See Willingham, 1988, *passim*.

[19]There are ways, however, to mitigate the costs. Assuming that one can, through special means, extend through some definable range the test coverage of the study, one may well be satisfied with but a subsample of these cases, and the use of weighting to ensure generalizability to the relevant population.

of the student questionnaire). Production of multilingual written materials must be supplemented by one-on-one oral administration to be truly effective. In the case of the student questionnaire, translation into (at least some) other languages is feasible, and although there is a question about possible mode of administration effects on the comparability of data collected, the document can indeed be readily administered by an interviewer, as is generally the case for documents designed for self-administration. Production of valid, truly parallel multilanguage test forms would be an enormous (and staggeringly expensive) undertaking. Moreover, in the case of the cognitive test battery, oral administration of the reading test is not possible, and is at least suspect for the other tests, grounded as they are in visual formats. To be usable, special testing arrangements must produce results which are validly and reliably linkable to the overall test battery results.

Recommendations with Respect to Levels of Inclusion. In many respects, it would seem preferable to view participation in national studies as admitting of degrees, rather than as a dichotomy centering on test-taking eligibility. To gather no data for students who cannot be tested infects national estimates with undercoverage bias and severely impedes the capacity of the study to examine such policy-relevant groups as LEPs/NEPs and physically and mentally handicapped students. Longitudinal studies that would create multiple nationally-representative grade cohorts are particularly obliged to follow all members of the cohort over time, regardless of baseline status. Consideration should therefore be given to the strategy of excluding no student from a study, but of possibly excluding that student from one or more of six possible tiers of participation, while including the student at all levels possible:

level 1 -- student enrollment status, demographics, detailed information about disabilities etc. from a school source

level 2 -- academic transcripts[20]

level 3 -- student questionnaire

level 4 -- test

level 5 -- contextual data questionnaires: ratings of student by teachers and supplemental teacher questions on student hadicapping conditions; parent questionnaire data; school questionnaire data on school characteristics, climate, practices, etc.

level 6 -- ecological data -- for example, student- and school-linked census tract data

Even if we take steps to reduce their numbers, there will remain some students who cannot participate at the fourth level, cognitive test completion, but can participate at level 3 (and can be represented at levels 1, 2, 5, and 6). Even when a student can complete neither test nor questionnaire, demographic and enrollment status, transcript, contextual and ecological data can be collected and reported. In

[20]Typically transcripts studies collect, for the student's entire high school career, such data as courses completed, credits earned, grades, days absent per year, participation in specialized programs (special education, bilingual education, gifted and talented education, and so on), class rank, class size, date and reason student left school (diploma, certificate of attendance, GED, dropped out, and so on), grade point average, available test scores (e.g., SAT, PSAT, ACT).

recent years, transcripts have been gathered for students excluded from testing (the NAEP 1987 and 1990 transcripts studies exemplify this, as does the 1992 NELS:88 high school transcripts study, which sought transcripts for eligible and ineligible high school seniors alike).[21] Likewise, the NAEP excluded student questionnaire, and the information gathered in the excluded student follow-backs of NELS:88, are an attempt to gather information at level 1 above. Nonetheless, studies such as HS&B and NELS:88 have not sought parent, teacher or other external data for students who have not completed the research instruments or who were deemed incapable of completing the instruments.

In addition, it is necessary to collect information permitting LEP and IEP students (whether capable or incapable of completing the survey forms) to be separately distinguished on the dataset, with meaningful detail (for example, categorization into the multihandicapped, mentally retarded, hard of hearing, deaf, speech impaired, visually handicapped, deaf/blind, seriously emotionally disturbed, orthopedically impaired, specific learning disabled, and other health impaired) provided about the nature of their disabling conditions.

Recommendations with respect to baseline transfer students. While the largest source of undercoverage in the NELS:88 base year is the 5.37 percent of the sample that was excluded from participation, an additional source of undercoverage is the under-representation of transfer students. An expensive and logistically complicating solution to the problem of transfers between the time of sampling (in HS&B and NELS:88, the fall term) and surveying (in HS&B and NELS:88, the spring term) of students would be to follow all students once they have been selected. Particularly from a school effects research perspective, the student who has just left a school may be of more interest than the one who has just transferred to a new school.

A more cost-efficient strategy, and one which maintains design simplicity, is to continue the strategy of excluding transfers-out and sampling from the transfers-in in a sample update just prior to the school's survey session, but to accommodate undercoverage of transfer students in the weighting. One should collect race/ethnicity, gender, and other basic information about the sample at the time of initial sampling. Weights for transfer-out students should be calculated. The estimated undercoverage of transfer-ins would be accounted for by modifying the weights of this group appropriately.

Conclusions. The experience of NELS:88 suggests that although it is not possible to include all students in self-administered surveys and assessments, more students have typically been excluded than is justified. Beyond sharpening definitions for exclusion and operationalizing them in a more systematic and effective way, some investment in special measures could also increase the number of students who are included in educational data bases. It is also possible to sample expensive "special effort" cases at a lower rate and make appropriate statistical adjustments on the basis of this added information, if cost constraints so dictate. Over and beyond increasing overall questionnaire and test coverage, it would also be prudent to include, by other means, those students who cannot complete research instruments. Their progress through school can be followed, based on school reports, their transcripts, and linkages to school and home contextual and community ecological data sources.

---

[21]While collecting transcripts for all students ensures the representativeness of estimates that come out of the dataset, to study handicapped students as a separate group, it is important that IEP information also be collected from the school for use in conjunction with transcripts data.

REFERENCES

Algozzine, B., and Ysseldyke, J. 1981. "Special Education Services for Normal Children: Better Safe Than Sorry?" **Exceptional Children**, 48.

Allen, T.E. (Ed.) 1989. "Handicapped Youth and the Outcomes of Their Schooling: Recent Research Advances in the United States." **International Journal of Educational Research**, 13.

Bennett, R.E., and Ragosta, M. 1984. **A Research Context for Studying Admissions Tests and Handicapped Populations**. Princeton, N.J.: Educational Testing Service.

Bennett, R.E., and Ragosta, M. 1988. "Handicapped People." In Willingham, et al.

Braun, H., Ragosta, M., and Kaplan, B. 1988. "Predictive Validity." In Willingham, et al.

Bruininks, R., Thurlow, M. L., and Ysseldyke, J. E. 1992. "Assessing the Right Outcomes: Prospects for Improving Education for Youth with Disabilities." **Education and Training in Mental Retardation**, 27.

Craig, P., Myers, E., and Wujek, M. 1982. **Independent Evaluation of the California Master Plan for Special Education: Final Report.** Menlo Park, CA: SRI.

GAO (U.S. General Accounting Office). 1993. **Student Testing: Current Extent and Expenditures, With Cost Estimates for a National Examination**. Washington, D.C.: GAO/PEMD-93-8.

Henderson, A., Daft, J., and Gutmann, B. 1989. **A Summary of State Chapter 1 Migrant Education Program Participation and Achievement Information, 1986-87**. Washington, D.C.: Decision Resources Corporation.

Hoachlander, E.G. 1992. **Participation in Secondary Vocational Education, 1982-87.** Washington, D.C.: National Center for Education Statistics.

Ingels, S.J., Rizzo., L. and Rasinski, K.A. 1989. **School, Individual, and Item Nonresponse in the National Education Longitudinal Study of 1988 (NELS:88) Base Year Survey**. Paper presented at the AERA annual meeting, San Francisco. ERIC ED 312-311.

Ingels, S. J. 1991. **The Problem of Excluded Baseline Students in a School-Based Longitudinal Study**. Paper Presented at the AERA annual meeting, Chicago. ERIC ED 334-273.

Ingels, S.J., Scott, L.A., Lindmark, J.T., Frankel, M.R., and Myers, S.L. 1992. **NELS:88 First Follow-Up Student Component Data File User's Manual**. Washington, D.C.: National Center for Education Statistics. (NCES 92-030.)

Johnson, E.G., and Zwick, R. 1990. **The NAEP 1988 Technical Report**. Report to NCES prepared by the Educational Testing Service, Princeton, N.J.

Linn, R.L. 1993. "Educational Assessment: Expanded Expectations and Challenges." **Educational Evaluation and Policy Analysis**. 15(1).

McGrew, K.S., Thurlow, M.L., Shriner, J.G. and Spiegel, A.N. 1992. **Inclusion of Students with Disabilities in National and State Data Collection Programs**. Minneapolis: National Center on Educational Outcomes.

National Council on Disability. 1993. **Serving the Nation's Students With Disabilities: Progress and Prospects**. Washington, D.C.: Author.

Owings, J.A., and Stocking, C. 1985. **Characteristics of High School Students Who Identify Themselves as Handicapped**. Washington, D.C.: National Center for Education Statistics.

Shepard, L.A., Smith, M., and Vojir, C. 1983. "Characteristics of Pupils Identified as Learning Disabled." **American Educational Research Journal**, 20.

Spencer, B.D., Frankel, M.R., Ingels, S.J., Rasinski, K.A., and Tourangeau, R. 1990. **NELS:88 Base Year Sample Design Report**. Washington, D.C.: National Center for Education Statistics.

Spencer, B.D. 1991a. Eligibility/Exclusion Issues in the 1990 Trial State Assessment. In G. Bohrnstedt et al., **Assessing Student Achievement in the States: Background Studies for the Evaluation of the NAEP Trial State Assessment**. New York: National Academy of Education.

Spencer, B.D. 1991b. A Critique of Sampling in the 1990 Trial State Assessment. In G. Bohrnstedt et al., **Assessing Student Achievement in the States: Background Studies for the Evaluation of the NAEP Trial State Assessment**. New York: National Academy of Education.

U.S. Bureau of the Census. 1993. **Language Spoken at Home and Ability to Speak English for United States, U.S. Regions and States: 1990**. Washington, D.C.: Author, CPH-L-133.

Willingham, W.W. 1988. "Discussion and Conclusions." in Willingham, et al.

Willingham, W.W., Ragosta, M., Bennett, R.E., Braun, H., Rock, D.A., and Powers, D.E. 1988. **Testing Handicapped People**. Needham Heights, MA: Allyn and Bacon.

Wolman, C., Bruininks, R.H., and Thurlow, M.L. 1989. "Dropouts and dropout programs: Implications for special education." **Remedial and Special Education**, 10.

Ysseldyke, J. 1983. "Current Practices in Making Psychoeducational Decisions About Learning Disabled Students." **Journal of Learning Disabilities**, 16.

Ysseldyke, J. Algozzine, B., and Epps, S. 1983. "A Logical and Empirical Analysis of Current Practices in Classifying Students as Handicapped." **Exceptional Children**. 50.